Reviewer’s report

Title: Elaborating on the Assessment of the Risk of Bias in Prognostic Studies in Pain Rehabilitation using QUIPS – Aspects of Interrater Agreement

Version: 0 Date: 07 Nov 2018

Reviewer: Danielle van der Windt

Reviewer's report:

This paper reports on the experiences of a systematic review team when using the QUIPS tool for assessing risk of bias in a systematic reviews of prognostic factors for long-term outcomes after multidisciplinary rehabilitation in patients with chronic pain. They report inter-observer agreement for risk of bias assessment across all domains and for each of the six risk of bias domains, and make suggestions for improving the process of scoring risk of bias of prognostic factor studies.

(1) Background: The introduction includes a paragraph on meta-analysis (estimating average effect size), which is fine but is not specific to prognosis reviews and not entirely relevant to the objective of this paper. It is also not quite clear if the paper by Page (reference 10) concerns prognosis studies or any type of study design. It would be good to focus the introduction on the importance of assessing quality and risk of bias in prognosis studies specifically. The weaknesses in prognosis research and the importance of assessing risk of bias in these types of studies have been highlighted in many papers (e.g. BMJ series by Moons & Altman; PROGRESS series; Hemingway et al. 2010; Hayden et al, 2013; Kyzas et al 2005; etc), and it would be helpful to highlight issues of specific importance to prognosis studies.

(2) Results / Discussion: I am not quite sure why results and discussion are written as one section - this could be easily separated?

(3) Results: inter-rater agreement - If I read this paragraph correctly, the results show that in 25/43 papers at least 4 domains were scored the same by the two observers. Describing agreement in a cumulative way might be more informative than describing agreement in 1, 2, 3, 4, 5, or 6 domains.

Results: Expected agreement is dependent on the distribution of the marginal totals (distribution of low, moderate, or high risk of bias) and consequently kappa is highly sensitive to this. Table 1, which presents the full 3x3 table shows that risk of bias was not evenly distributed and numbers are low for some categories. This may be more prominent for the analyses of agreement per bias domain, where numbers are smaller. This needs to be clear in the results: perhaps add a
table to present information regarding the marginal totals (e.g. for each domain, the proportion (n/N) assessed as high/moderate/low/risk of bias for each observer).

(4) Discussion: I am wondering if some of the findings for agreement, especially by risk domain, are influenced by the dependence of kappa on the distribution of margin total (see previous comment), and this might partly explain the discrepancy noted by the authors compared to other studies investigating the reliability of the QUIPS tool. On page 12, the authors do discuss the fact that agreement seemed higher in reliability studies where a larger proportion of papers included in the analysis were of better quality - which might indicate a more balanced distribution. Although there has been some debate in the literature to what extent this dependence of kappa on margin totals (or: prevalence) is a problem and how this should be dealt with, I think this needs to be acknowledged in the discussion.

(5) Discussion: The authors do not appear to report on the strengths or weaknesses of their study, e.g. was the number of papers large enough for the analysis? Are the results generalisable to other reviews conducted in other clinical areas, number of observers, …?

(6) Conclusions: The authors indicate that they put forward suggestions for improving the tool. I am not sure if they have. They elaborate on the process of using QUIPS, and how this can be improved (e.g. by providing more detailed guidance for scoring each signalling items, and how to weigh the different items when assessing risk of bias; by making sure observers are sufficiently knowledgeable; by embedding discussions regarding interpretation of items). This makes perfect sense and is important (tends to be important when assessing risk of bias for other types of studies as well), but do they have recommendations for amending the tool itself? And if so, what would be the next steps for improving and re-testing the tool?

(7) Typo: page 6, line 18: "included presented".

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I was a co-author on the paper by Hayden et al. introducing the QUIPS tool.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.